Market–based Decentralized Profile Infrastructure: Giving Back to the User

Olfa Nasraoui

Dept. of Computer Engineering & Computer Science
University of Louisville
Who Owns **YOUR** User Profile...?

- **Your Usage Profile:** Your historic data: pages viewed, items purchased, clickstreams, ...etc

- A loyal consumer/user of a Web business still cannot *own* their profile

- and cannot move it with them freely from one business (website) or context to another

- However this profile = precious information:
  - can be used for personalization
  - can combat information overload

Olfa Nasraoui, Market-based Decentralized Profile Infrastructure: Giving Back to the User
Different Levels of Correlations

- Likely correlations between a user's tastes in books, movies, and many other products or content items that are not sold on the same website
  - including: food, clothing, "content" like "news and blogs", music, ...etc
- need single profile integration across multiple websites
- Above correlations can only be enriched if further integrated with many other user profiles in a collaborative filtering (CF) framework,
  - CF: predicting a user's interests not just from the same user but also from other similar users' interests "by association".
- need multiple profile integration
Two **Scopes: User & Website**

- Currently **each website** has a limited view of user profiles (**limited** to what is sold/served by website).
- Extending the scope to **other websites** would give a more global view of a user profile.
- Currently **each user's scope** is limited: only their own profile is available, hence no sharing with other users.
- ➔ **a user cannot possibly be the one who "invokes" a collaborative filtering recommendation**
- Instead it is **always the server that initiates and benefits** from such collaborative filtering.
Proposed: Intermediate Solution

- Intermediate solution fosters both:
  - single profile integration (single profile across multiple websites) and
  - multiple profile integration (across multiple profiles on the same website).

- Midway between the server (or the business) and the client (or the user).

- Similar to peer to peer information sharing:
  - there is no single central control of the user profiles,
  - though there could be several repositories of many user profiles in server communities or clustered repositories.
Who Owns the Profile?

- A user, who **owns** his/her own profile, **earns some credit** each time that their profile is invoked by a recommendation process (or transaction),

- The individual credits may be very small, but may **accumulate** to a profitable level with a large number of invocations

- The user, for the first time, **not only "owns" their own profile, but also "can sell it like a commodity"** and benefit from it
Physical Infrastructure: P2P & Web Services

- Peer to peer networks: Information is exchanged in a decentralized manner
- Web services platforms: self-contained, self-describing, modular applications that can be published, located, and invoked across the Web
  - Services are published via a registration mechanism in a registry, and
  - Service requestors find Web services by checking these registries
  - Registry description (in Web Services Description Language, WSDL) contains sufficient information for the service requestor to bind to the service provider to use the service.
  - One big difference compared to user profile marketplace: the user does not necessarily have a "server" or a unique URL where they can be reached
Physical Infrastructure: **User’s Personal Mobile Device**

- Can use user’s mobile device/phone/computer
- Intermediate architecture:
  1. user information is **logged on their device in real time**
  2. then transmitted to one of several repositories that could be on actual servers/registries,
     - but **not necessarily a central server**
     - Example: "super peers“ in P2P networks: designated peers that act in a role halfway between a true server and a true peer in the exchange and routing of information

Olfa Nasraoui, Market-based Decentralized Profile Infrastructure: Giving Back to the User
Physical Infrastructure: **Social Networking Websites & RFID**

- **Social networking websites**: another vital “carrier” and platform for such an infrastructure
  - if a user profile can be stored and invoked securely as part of a market strategy
- **RFID readers implanted on the user's personal device, such as a cell phone** can integrate even offline transactions with online transactions
  - **Client computer**: Online transactions can be logged entirely by software
  - **RFID-equipped cell phone**: communicates this information with a designated repository
- ➔ **RFID can bridge the gap between online and offline worlds**

Olfa Nasraoui, Market-based Decentralized Profile Infrastructure: Giving Back to the User
Physical Infrastructure: e–wallets

- Integrating consumer profiles together with their e–wallets
- E–wallets: recently implemented based on the existing “mondex” card:
  - shares the flexibility and privacy features of cash
  - while allowing participation in electronic commerce even at micro–economical level (very small value transactions) (Tam & Ho, JOCEC’07)
Privacy

- **Taxonomy** of what can be logged or not logged among the user's interactions and purchases
  - Example: some users do not like to log anything that is related to health or finance
  - These restrictions will be directly implemented on the user’s “data collection” side

- **Privacy in Peer-to-Peer: 2 categories:**
  - **User anonymity:** hide identity of user requesting info
    - Existing frameworks: e.g. Freenet
  - **Data privacy:** protect sensitive info
    - Privacy preserving data mining: discover useful patterns without compromising privacy (e.g. while keeping the input data private/hidden)
    - E.g.: Privacy preserving *similarity computation* in P2P networks: there are existing methods that can compute an output while keeping the input data private/hidden) (e.g. Goethals et al., ICISC’04; Liu et al, WebKDD’06)
    - as by-product: also help support a payment mechanism (*per invocation*)
Possible Interaction Models: Markets

- **Item sellers**: companies that traditionally sell products
- **Item buyers**: customers who have made prior transactions with these or other sellers.
- **A seller is interested in computing good recommendations** for items that they sell based on the profiles of available buyers with matching profiles.
- **A recommendation has a value to the seller** if it improves the seller’s recommendation model or if it results in a sale.
- **Buyers are interested in offering (selling) their profile information** in return for a reward from the seller.
- **Seller & buyer’s optimization problems**: can cast as primal and dual complementary linear programs.
Possible Interaction Models: Graphs & Social Networks

- **user–item graph:**
  - **Nodes** = users and items
  - **Edges** = connect a user and an item if this item belongs to the user’s profile
  - **User–item links** = probabilities of following a certain user–item link
  - **Cosine similarity** between two users (in a user–based Collaborative Filtering) can be viewed as computing the probability that the two users will ever meet at any location while doing a random walk on the above graph

- **user–user graph:**
  - **Nodes** = users
  - **Edges** = weighted by cosine similarity between two user profiles
  - or **Edges** = strength of connection in a social network.

- **item–item graph:**
  - **Nodes** = items
  - **Edges** = connect 2 items if they occur in the same transaction at any time
Possible Interaction Models: Query Incentive Networks

- A random walk on a graph model = basis for many Collaborative Filtering (CF) recommendation strategies

- Recommendation process takes place by:
  - submitting a query to one or more nodes in the graph,
  - allowing these nodes to pass this query on to their local neighbors,
  - and then waiting for an answer which is returned when a satisfactory answer is found.

- Similar to information retrieval in P2P network

- However, without an incentive for users to participate, the effective active (i.e. responding) network at any time could be very limited

(Kleinberg and Raghavan, FOCS '05)
Possible Interaction Models: Query Incentive Networks

- Rather than posing queries to a *centralized index* of the system, users pose queries to the *network* itself.
- Requests for information are propagated along paths through the network, connecting those with information needs to those with relevant answers.
- Queries are submitted together with incentives for answering them.
- Incentives get propagated along paths in a network,
- Each participating node earning a portion of the reward,
- Until either an answer is found or the propagating rewards get depleted...
- This information-seeking process was formulated as a game among the nodes in the network.
- Game was shown to possess a natural Nash equilibrium.
Possible Interaction Models: Another game theoretic approach

- Implement infrastructure by means of dynamic and real-time *automated auctions* between companies and consumers

- **Equilibria** of such systems: can shed light on its promises, as well as whether it provides fair play for all the players, and under what conditions
Graph-based Interaction Models: Challenges

- **May not scale** to millions of transactions per second
- **Index based retrieval**: may be the only option on several indexed Databases,
  - but the DBs need to be refreshed with every user’s new transactions

- **Impractical** to perform graph search for **each transaction**
  - Perform **searches in bundles of transactions or users**: i.e. find the optimal profiles for a batch of transactions
  - Perform **searches on sub-graphs**: E.g. after offline discovery of communities within the original graph
Graph–based Interaction Models: Advantages

- Graph model supports a **distributed** profile base,
  - no single authority owns all profiles

- Graph–based search supports **local** search:
  - where a query is passed from one node to its neighbors,
  - thus limiting threats to privacy
Bio-inspired Models

- **Co-evolutionary models**: two or more species co-evolve simultaneously to optimize their survival.
- **Ant Colonies**: a community of simple agents (ants) can succeed well in achieving collaborative tasks, using a set of simple rules governing each individual ant,
  - further enriched by communication ➔ special collaborative intelligence known as *stigmergy*
- **Swarms**: agents (e.g. bees) move in such a way that each agent is aware of the movement and fitness of its neighbors,
  - ➔ resulting in complex behavior.
Conclusions

- **Market-based profile infrastructure:**
  - Put the users more in control of their own profiles,
  - Allow the user to own and profit from this profile, thus democratizing recommendations

- **Several challenges, research issues and possible solutions**: far from complete

- **Monetary system**: e-wallets such as mondex or paypal accounts?

- **Physical infrastructure**: existing P2P networks, social networking websites, Web services infrastructure

- **Theoretical & Empirical analysis and modeling**: Study and simulate the market dynamics, Optimization: of costs & profits, Graph models, for community discovery, Query incentive networks, Social networks

- **Ontological engineering**: to relate items in profile ➔ Need category information (taxonomy) or textual description of an item

- **Privacy Preservation**: What are the risks? What are the remedies?
  - Can the reward outweigh the risk? Sometimes?

- **Ethical & Legal Issues**: Who wins, what is fair? who stores and controls the data?

- **Other Issues? Solutions?** It’s your turn for ideas…

Olfa Nasraoui, Market-based Decentralized Profile Infrastructure: Giving Back to the User